U.S. ENVIRONMENTAL PROTECTION AGENCY, REGION 5

ELECTRONIC RECORDS MANAGEMENT POLICY AND DIGITIZATION REQUIREMENTS

U.S. ENVIRONMENTAL PROTECTION AGENCY, REGION 5, ELECTRONIC RECORDS MANAGEMENT AND RECORDKEEPING POLICY AND DIGITIZATION REQUIREMENTS

Approval for Implementation:

DAVID HOFF
Digitally signed by DAVID HOFF
Date: 2021.03.11 14:23:27 -06'00'

David Hoff, Region 5 Records Liaison Officer — Date

Carmen Maso
Digitally signed by Carmen Maso
Date: 2021.03.11 14:25:57 -06'00'

Carmen Masó, Section Chief, Information Services Section — Date

KENNETH TINDALL
Digitally signed by KENNETH TINDALL
Date: 2021.03.11 14:29:12 -06'00'

Kenneth Tindall, Branch Chief, Information Management Branch — Date

AMY SANDERS
Digitally signed by AMY SANDERS
Date: 2021.03.11 14:37:05 -06'00'

Amy Sanders, Mission Support Division Director — Date

CHERYL NEWTON
Digitally signed by CHERYL NEWTON
Date: 2021.03.15 14:20:02 -05'00'

Cheryl Newton, Deputy Regional Administrator — Date

U.S. ENVIRONMENTAL PROTECTION AGENCY, REGION 5, ELECTRONIC RECORDS MANAGEMENT AND RECORDKEEPING POLICY AND DIGITIZATION REQUIREMENTS

March 11, 2021

TABLE OF CONTENTS

I. INTRODUCTION

This document establishes requirements for electronic record keeping throughout the U.S. Environmental Protection Agency (U.S. EPA), Region 5 (R5). Since 2011, records management within U.S. EPA and other agencies in the executive branch have been increasingly drawn away from traditional paper records management and toward electronic methods of records preservation and retrieval. Section 2.4 of the National Archives and Records Administration (NARA) Strategic plan states, "By FY 2020, NARA will have policies and processes in place to support Federal agencies' transition to fully electronic recordkeeping." After December 31, 2022, to the fullest extent possible, both Federal Records Centers (FRC) and NARA will no longer accept permanent and temporary records transfers in paper format. The U.S. EPA Agency Records Officer intends to comply with this requirement and does not foresee any exceptions. The R5 Document Management Team has observed an increasing business need at the regional level to respond in a transformative way to these changes and bring our recordkeeping practices into compliance with NARA's overall strategic plan. The necessary changes in regional business practices are all possible with expansion and adoption of existing systems. This will position the region to be able to respond efficiently and flexibly to an ever-changing electronic records landscape. The default policy for the region will be to prefer the original, electronic version of any record. The electronic record is to be preserved in one of the available records systems, and any paper copies are to be considered convenience copies of the original record.

As of the date of this policy, Region 5 will no longer consider external drives suitable for records preservation or storage. Any electronic records created or stored on external drives must be properly added to one of the existing records systems. To align with regional policy on text message or personal email, users will have 20 days to preserve any electronic records in an appropriate system. Additionally, procurement requests for external storage devices will no longer be approved by the Region 5 IMB Chief/R5 Senior IT Leader (SITL) to reinforce the use of Agency, secure cloud storage services.

II. Electronic Records Systems

Region 5 currently uses four electronic records systems. The two national systems are the Superfund Enterprise Management System (SEMS) and the Electronic Content Management System (ECMS) which is the destination system of EZ Records. The two Regional Systems, both built in SharePoint, are the Electronic Document Management System (EDMS) and the Electronic Records Center (ERC). Most Regional programs have adopted one or more of the existing systems, and a concerted effort is being made to expand and configure the existing systems to accommodate all regional records.

The following items detail proposed systemic enhancements or changes to each system, appropriate to these objectives:

- **SEMS-RM:** Currently, R5 is transferring SEMD, LCRD/RCRA, and GLNPO/GLLA site-specific information assets into SEMS-RM which, in the case of paper, proceed to the FRC after digitization and quality review. Recent centrally implemented enhancements within SEMS-RM include the ability to establish "Non-CERCLA Site" listings at the System Administrator level. The functionality has already been successfully demonstrated for non-Superfund facilities by adding Areas of Concern under the GLLA to transfer information assets documenting remedial actions conducted by GLNPO in these areas. There is no functional restriction within SEMS-RM to prevent inclusion of other geographic entities such as facilities regulated under any other Non-CERCLA site category. As of publication of this policy, WD/UIC/Class V permits and ORA/NEPA files cradle to grave are also in SEMS.

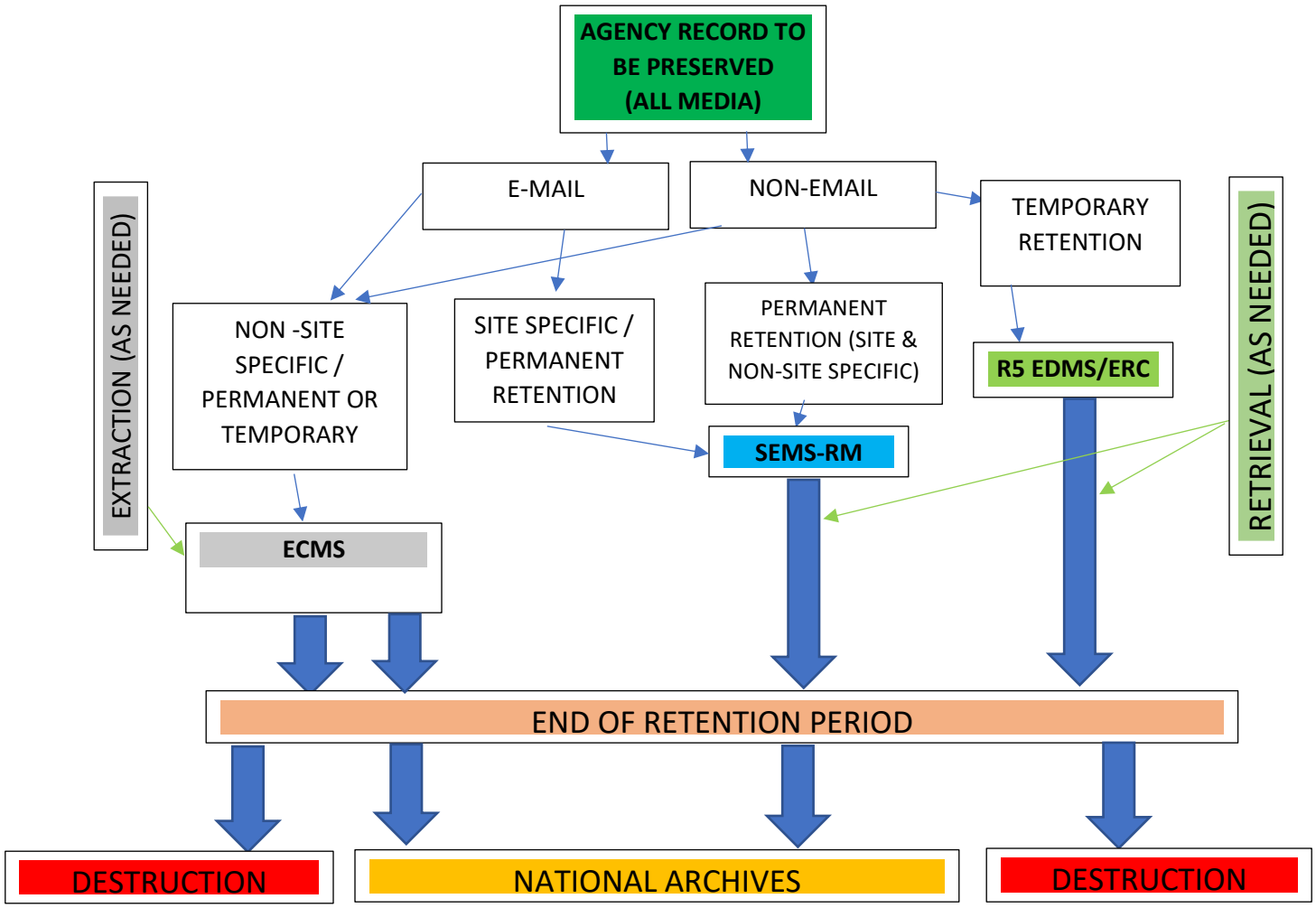
- Each Records Officer would serve as a SEMS-RM System Administrator with the ability to add Non-CERCLA names on an as-needed basis. Meta-tagging will be standardized, augmented, and customized to include program-specific identifiers for divisions other than SEMD, LCRD, and GLNPO. The Records Management task orders in R5 will develop the augmented multi-divisional metatags and incorporate the procedures and taxonomy into the R5 SEMS-RM Production Manual. In addition, cross-training in SEMS-RM indexing so that the Regional records contractors will, in the medium term, begin digitizing Non-CERCLA records into the system.
    - Site Specific email would be preserved using the Outlook to SEMS-RM connector (in development). As an interim solution, E-Mail can be submitted into SEMS using the RTI module, or forwarded to R5_SFRecords@epa.gov (low volume). Site-specific E-Mails can also be extracted from ECMS/EZ-Records for preservation in SEMS-RM.

- **R5 EDMS:** EDMS is being phased out and existing records in the system are being migrated to ERC. The R5 EDMS will primarily be maintained as a legacy system following full implementation of the R5 ERC.

- **<u>ECMS / DOCUMENTUM</u>:** Will continue to serve as the "EZ Records/Desktop" preservation system for **non-site-specific** emails and non-emails, with SEMS-RM and R5 ERC serving the programs in a retrieval capacity. EZ Records can also serve as an ad-hoc tool for quick preservation in "as needed" scenarios, such as short-notice agency separation by departing staff. Preserved ECMS e-mails can be extracted in response to future programmatic needs.

- **R5 ERC:** Is a repository and retrieval system for temporary, non-email, non-site-specific records. Flexibility for program/division use, i.e., "policy will be to work with divisions to develop processes within ERC or their own SharePoint repositories, to ensure proper tools are in place to: 1) associate records with their retention schedules; and 2) monitor records retention activities." When changes/updates are made to records in ERC, the new version will be placed in ECMS. The ERC is the only existing electronic records system that includes records scheduling and applied retention periods. ERC has been adopted by several programs and is being configured to accommodate the rest of the region.

PROCESS FLOWCHART FOR REGION 5 RECORDKEEPING UNDER THE NEW PROPOSAL



III. REGIONAL PRODUCTION STANDARDS

Purpose

An integral component to achieving regional transition to electronic records management includes verification that the capture of an electronic record into an official agency recordkeeping system meets a sufficient level of quality to allow the digital product to supplant any originating paper record. The purpose of this guidance document is to identify a minimum processing level of effort at the regional level to ensure that a digital rendering of a legacy paper record is a true and searchable electronic record equivalent to the source material's visual quality, and thus selectable by R5's Records Management Program as the preferential record over the hard copy. This guidance portion of this document is separated into two primary sections: a minimum rendering standard and a quality management process.

Effective May 10, 2019, NARA adopted new standards for digitizing temporary records. Section 1236.2 of the general NARA regulations were modified to clarify the terms, Digitization, Electronic Information System, Electronic Mail System, Metadata and Unstructured Data. Future amendments to this regional guidance addressing metadata, mixed-media, digitization, and production will ensue as the NARA's digitization standards for temporary and permanent records continue to be implemented.

Scope

Region 5 currently uses two primary document management systems to store digitized renderings of hard copy records - the regional SharePoint-based R5 ERC, and the national SEMS-RM for site-specific records. The regional standards are intended to apply to the regionally designated records management systems for temporary and permanent record capture, retention, and dispositioning.

Minimum Rendering Standard

All digitized output in R5 originating from a paper source shall be rendered to a minimum 300 PPI for textual documents, and 600 PPI for graphical information assets. Higher resolutions may be applied, as needed, to capture finer details. In addition, all applicable (textual) production output should be processed via Optical Character Recognition (OCR), with the resulting product full-text searchable.

- Standard file format options for production output digitized, captured, encoded, or otherwise retained shall be as follows:

Format Name and Version	Acceptable Lossless Compression Codecs
TIFF 6.0	Uncompressed, LZW compression
JPEG2000 part 2 (Annex M - JPX baseline)	JPEG 2000 part 1 core coding system lossless compression
Portable network graphics 1.2 (PNG)	DEFLATE (ZIP)
PDF/A-1	DEFLATE (ZIP)
PDF/A-2	DEFLATE, JPEG 2000 part 1 core coding system lossless compression

- Standard resolution specifications for color graphical materials digitized, encoded, or otherwise retained shall be as follows:

Digital file specifications[1]	Attributes
Color mode	RGB color or grayscale
Bit depth	8- or 16-bit
Color space	gray gamma 2.2, AdobeRGB1998, sRGB, ProPhoto, ECIRGBv2
Spatial resolution	600 PPI minimum

Quality Management Process (QM)

QM is the overarching program ensuring regional production output integrity, determining quality policies, objectives, and responsibilities, and implementing them through planning, control, assurance, and improvement methods within the production standards. QA is a series of steps taken in advance of digitization to prevent defects. QC is a series of procedures to inspect production output after digitization has taken place, and to verify that the digitized product is suitable for long-term preservation in an agency recordkeeping system. The combined steps comprise an effort to ensure that the digitized product will meet the

regional and agency specifications, particularly integrity of hard copy digitization, as defined under CIO 2155-S-01.0. Practical steps as listed will be correlated with specific components of both CIO 2155-S-01.0 and CIO 2155-P-05.0 to serve as standard QA/QC compliance process.

The following steps are to be taken in R5 to ensure that any digitized product complies with the agency digitization standards and procedures, as identified in the previous paragraph, NARA regulations for Digitizing Temporary Federal Records-Subpart D—(and forthcoming permanent records digitization standard). The steps are organized into QA and QC, as identified in the QM process introduction. Where relevant, the correlating section in the standards document will be cited in parenthesis. These steps apply to both SEMS-RM and R5-EDMS unless otherwise notated to specify the applicable records management system.

Quality Assurance (QA)

Steps in advance of scanning operations, closing a batch, or releasing documents:

- Equipment shall be regularly tested to ensure scanners and digital cameras/copy systems are performing optimally.
    - (i) Reference target shall be scanned containing a grayscale, color chart, and accurate dimensional scale at the beginning of each workday; and
    - (ii) Perform additional tests when problems are detected;
    - (iii) Monitor device performance during digitization; and
    - (iv) Verify that resulting digital files meet project specifications.
- Ensure the integrity of the originating hard copies by doing the following:
    - Keep incoming documents sequenced in the same order as received.
    - If digitizing from an archival accession, check to make sure that folders listed on the box inventory match what is in the box. Note: need to update box lists and/or metadata with corrections during the QC process. (CIO 2155-P-05.0 – Section 6/F).
- Be on the lookout for documents that do not require scanning, e.g., a "DNS" (Do Not Scan for SEMS-RM), or another relevant notation from the indexer.
- Be sure to scan all Post-its® or "sticky" notes *left attached* by the indexer with substantive content from the original submitter, and ensure the attached notes are not covering any text on the document.
- Verify which oversized materials will require a wide format scanner (beyond 11x17), such as large maps, boards, thick folders, or binders where the use of the duplex scanner is not applicable. If this functionality is not available, insert an "Unscannable Page" that gives an explanation as to why a page is unscannable. In such instances, the hard copy needs to be maintained. (CIO 2155-P-05.0 – Section 6/F).

Quality Control (QC)

Steps taken by the scanning operator during scanning:

- Check for validation (currently SEMS-RM) – Make sure scanning software is properly identifying every barcode tag. If the barcode does not register, a document can be manually assigned a document ID number. Check the Batch Manager to see whether all the documents are generated. Sometimes the scanner will separate pages even though there is no Doc ID.

Steps taken by the scanning operator after scanning:

- Check for image legibility – Assure that the documents are legible to the end user. Some pages of the documents may vary in resolution (PPI), appearing either very dark or too light and need to be adjusted to be made legible or cohesive with the other pages. Ensure all textual pages are at least 300 PPI. (CIO 2155-S-01.0 – Section 6/J), while graphical pages are digitized at 600 PPI.
- Verify that the document is searchable via OCR - For textual documents, select a minimum of three substantive terms that are known to exist within the document (ex. can be seen in the images) and run a CTRL+F search function to ensure that the terms can be retrieved. (CIO 2155-P-05.0 – Section 6/H)
- Check for page orientation – Verify whether the document is in the landscape or portrait position appropriately. Rotate pages if applicable. Rescan a page if it is crooked or skewed beyond 3 degrees of the hard copy's positioning. (CIO 2155-P-05.0 – Section 6/H).
- Check for zero images (currently SEMS-RM) – This is the last step when documents are completely scanned. The scanning operator should check the doc id range to ensure that all documents are completely scanned before they are dispositioned. This will avoid the tedious process of retrieving a document if it is not scanned properly.
- If you are using a dedicated scanner instead of the multi-use machines, clean scanning equipment periodically. Replace roller and other scanner supplies, as needed.

Post QA/QC

- Documents stored at the FRC should be returned in the same order they are received.

Hard copies that have been digitized and vetted in accordance with the procedures in this guidance document may be discarded once the equivalent digital version is known to exist in an agency recordkeeping system (SEMS-RM or R5-EDMS).

Glossary

CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act of 1980
CIO	Chief Information Officer (Signatory of Records Policies)
Documentum	Software for the Agency Electronic Recordkeeping System
ECMS	Electronic Content Management System (Agency Electronic Recordkeeping System)
EDMS	Electronic Document Management System (SharePoint System)
ERC	Electronic Records Center (SharePoint System)
FRC	Federal Records Center
GLLA	Great Lakes Legacy Act
GLNPO	Great Lakes National Program Office
IMB	Information Management Branch
ISS	Information Services Section
LCRD	Land, Chemicals & Redevelopment Division
MSD	Mission Support Division
NARA	National Archives & Records Administration
NEPA	National Environmental Policy Act
OCR	Optical Character Recognition
PDF	Portable Document Format
PDF/A#	Portable Document Format Archival Version
PPI	Pixels Per Inch
QA	Quality Assurance
QC	Quality Control
QM	Quality Management
RCRA	Resource Conservation and Recovery Act of 1976
R5	Region 5
SEMD	Superfund & Emergency Management Division
SEMS	Superfund Enterprise Management System
SEMS-RM	Superfund Enterprise Management System-Records Management
SITL	Senior IT Leader
U.S. EPA	U.S. Environmental Protection Agency
WD	Water Division